日本音響学会誌 71 巻 11 号（2015），pp. 607–614

小特集—スパース表現に基づく音響信号処理—

スパース性に基づく音楽音響信号の分解*

吉 井 和 佳，糸 山 克 寿（京都大学大学院情報学研究科知能情報学専攻）**

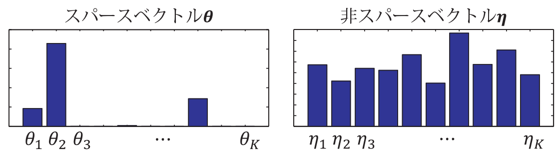
43.75.Zz; 43.60.Uv

1. は じ め に

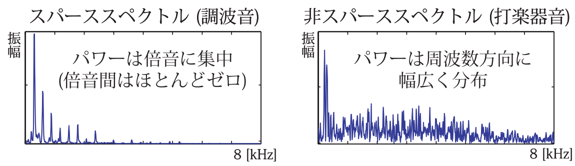
本稿では，スパース性に着目することで，音楽音響信号を構成要素に分解する技術について解説する。ベクトルや行列がスパースであるとは，ほとんどの要素がゼロをとる状態を指す（図–1）。一見複雑に思える音楽音響信号も，高々有限個の楽器音が重なりあってできている。例えば，あるピアノ曲であれば，出現する音高は音域や調に依存することから，88 鍵の使われ易さにスパース性が存在する。一方，音楽音響信号の構成要素にもスパース性が存在する。例えば，調波構造を持つ楽器音は倍音周波数付近にパワーが集中しており，倍音間にはほとんどパワーが存在しない（図–2）。

このような特徴を持つ音響信号のスペクトログラムを少数のスパーススペクトル（基底と呼ぶ）のスパースな線形和に分解するには非負値行列分解（Nonnegative Matrix Factorization: NMF）[1, 2] が有用である。音響信号全体では K 個の基底スペクトルが必要であるとしても，各時間フレームでは限られた少数の基底の線形和で表現可能である。この考え方を進めると，入力音響信号に合わせて基底数 K を手動で調整する代わりに，可算無限個の基底の存在を仮定し，必要な基底だけが自動的に実体化できれば好都合である。このような無限次元の空間（$K \to \infty$）におけるスパースな学習は，ノンパラメトリックベイズモデルを用いて実現することができる [3]。近年，NMF の数学的に自然な拡張である半正定値テンソル分解（Positive Semidefinite Tensor Factorization: PSDTF）[4, 5] が提案され，優れた音源分離結果を達成している。

また，構成要素のスパース性に着目することで，



図–1　K 次元のスパースベクトル $\boldsymbol{\theta}$ と非スパースベクトル $\boldsymbol{\eta}$：各ベクトルは各次元 k $(1 \le k \le K)$ に対して $\theta_k \sim \mathrm{Gamma}(0.1, 0.1)$，$\eta_k \sim \mathrm{Gamma}(10, 10)$ として生成。$\mathbb{E}[\theta_k] = \mathbb{E}[\eta_k] = 1$ であることに注意。



図–2　調波構造を持つ楽器音のスペクトル（ピアノ音）と打楽器音のスペクトル（スネアドラム）。

市販 CD のように複雑な音楽音響信号に対しても，歌声・伴奏音を分離したり，調波音・打楽器音を分離することができる。音色や音高のバリエーションに富む歌声を含む音楽音響信号は，NMF のように少数の基底の線形和で表現することは難しい。一方，ロバスト主成分分析（Robust Principal Component Analysis: RPCA）[6] を用いると，線形和で表現可能な低ランクな伴奏音成分と，外れ値と見なせるスパースな歌声成分とに分離可能である。また，時間・周波数平面において，調波構造を持つ楽器音は周波数方向にスパースであり，減衰が早い打楽器音は時間方向にスパースである。従って，周波数・時間方向にそれぞれメディアンフィルタを適用すれば，外れ値となる調波音・打楽器音をそれぞれ抑制することができる [7]。

2. 非負値行列分解

本章では，NMF に基づくモノーラル音響信号の音源分離について述べる。まず，コスト関数最小化としての定式化（最尤推定）と音源分離への適用について説明し，発展的内容としてノンパラメトリックベイズモデルの学習について解説する。

---

* Music signal decomposition based on sparseness.

** Kazuyoshi Yoshii and Katsutoshi Itoyama (Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University, Kyoto, 606–8501) e-mail: yoshii@kuis.kyoto-u.ac.jp

Algorithm 1 KL-NMF の最尤推定

Require: 非負値行列 $\boldsymbol{X} \in \mathbb{R}_+^{M\times N}$, 基底数 K
非負値行列 $\boldsymbol{W} \in \mathbb{R}_+^{M\times K}$ をランダムに初期化
非負値行列 $\boldsymbol{H} \in \mathbb{R}_+^{M\times K}$ をランダムに初期化
while not converged **do**
$\quad w_{km} \leftarrow w_{km} \frac{\sum_n x_{nm} h_{kn}/y_{nm}}{\sum_n h_{kn}}$
$\quad h_{kn} \leftarrow h_{kn} \frac{\sum_m x_{nm} w_{km}/y_{nm}}{\sum_m w_{km}}$
end while
return 非負値行列 $\boldsymbol{W}$, $\boldsymbol{H}$

Algorithm 2 IS-NMF の最尤推定

Require: 非負値行列 $\boldsymbol{X} \in \mathbb{R}_+^{M\times N}$, 基底数 K
非負値行列 $\boldsymbol{W} \in \mathbb{R}_+^{M\times K}$ をランダムに初期化
非負値行列 $\boldsymbol{H} \in \mathbb{R}_+^{M\times K}$ をランダムに初期化
while not converged **do**
$\quad w_{km} \leftarrow w_{km} \left(\frac{\sum_n x_{nm} h_{kn}/y_{nm}^2}{\sum_n h_{kn}/y_{nm}}\right)^{\frac{1}{2}}$
$\quad h_{kn} \leftarrow h_{kn} \left(\frac{\sum_m x_{nm} w_{km}/y_{nm}^2}{\sum_m w_{km}/y_{nm}}\right)^{\frac{1}{2}}$
end while
return 非負値行列 $\boldsymbol{W}$, $\boldsymbol{H}$

2.1 コスト関数最小化としての定式化

NMF では，入力となる非負値行列 $\boldsymbol{X} = [\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N] \in \mathbb{R}_+^{M\times N}$ に対し，

$$\boldsymbol{X} \approx \boldsymbol{W}\boldsymbol{H} \stackrel{\text{def}}{=} \boldsymbol{Y} \tag{1}$$

となる二つの非負値行列 $\boldsymbol{W} = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_K] \in \mathbb{R}_+^{M\times K}$，$\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K]^{\mathrm{T}} \in \mathbb{R}_+^{K\times N}$ への低ランク分解を行う。ただし，$\boldsymbol{w}_k \in \mathbb{R}_+^M$ 及び $\boldsymbol{h}_k \in \mathbb{R}_+^N$ はそれぞれ基底ベクトル及び対応するアクティベーションベクトルであり，基底数は $K \ll \min(M, N)$ とする。ここで，再構成行列を $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_N] \in \mathbb{R}_+^{M\times N}$ とすると，

$$\boldsymbol{x}_n \approx \sum_{k=1}^{K} h_{kn}\boldsymbol{w}_k \stackrel{\text{def}}{=} \boldsymbol{y}_n \tag{2}$$

と書ける。観測ベクトル $\boldsymbol{x}_n$ と再構成ベクトル $\boldsymbol{y}_n$ との間の誤差 $\mathcal{D}(\boldsymbol{x}_n|\boldsymbol{y}_n)$ を評価する尺度として，本稿では以下で定義される Kullback-Leibler（KL）ダイバージェンス [8] 及び Itakura-Saito（IS）ダイバージェンス [2] に着目する。

$$\mathcal{D}_{\text{KL}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_{m=1}^{M}\left(x_{nm}\log\frac{x_{nm}}{y_{nm}} - x_{nm} + y_{nm}\right) \tag{3}$$

$$\mathcal{D}_{\text{IS}}(\boldsymbol{x}_n|\boldsymbol{y}_n) = \sum_{m=1}^{M}\left(\frac{x_{nm}}{y_{nm}} - \log\frac{x_{nm}}{y_{nm}} - 1\right) \tag{4}$$

これらの値は常に非負であり，$\boldsymbol{x}_n = \boldsymbol{y}_n$ のときのみ 0 となる。また，通常の距離尺度と異なり，非対称性 $\mathcal{D}(\boldsymbol{x}_n|\boldsymbol{y}_n) \neq \mathcal{D}(\boldsymbol{y}_n|\boldsymbol{x}_n)$ に注意する。

2.2 乗法更新アルゴリズムに基づく最適化

コスト関数 $\mathcal{D}(\boldsymbol{X}|\boldsymbol{Y}) = \sum_n \mathcal{D}(\boldsymbol{x}_n|\boldsymbol{y}_n)$ を最小化する $\boldsymbol{W}$ 及び $\boldsymbol{H}$ を求めるため，乗法更新アルゴリズム [9] と呼ばれる反復最適化技法が利用できる。この手法は，各パラメータに対してある係数をかけて更新を行うため，係数が非負であればパラメータの非負性は自然に保たれる特徴がある。

本節では，補助関数法に基づく収束性が保証された乗法更新アルゴリズムを紹介する。KL-NMF 及び IS-NMF における更新則はそれぞれ **Algorithm 1** 及び **Algorithm 2** で与えられる（導出は文献 [9] 参照）。ただし，スケールの任意性を解消するため，$\sum_m w_{km} = 1$ を満たすよう，反復ごとに $\boldsymbol{w}_k$ 及び $\boldsymbol{h}_k$ をスケーリングしておく。

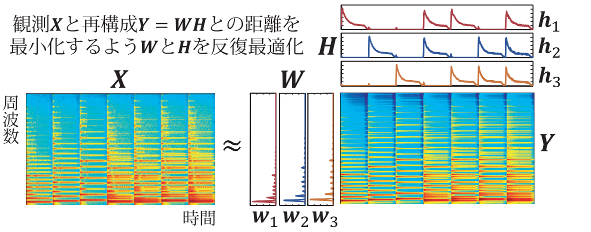
2.3 非負性に基づくスパースな分解

NMF では，基底ベクトル $\boldsymbol{w}_k$ 及びアクティベーションベクトル $\boldsymbol{h}_k$ がスパースになり易い。一方，同様の行列分解形式 $\boldsymbol{X} = \boldsymbol{W}\boldsymbol{H}$ を持つ主成分分析（Principal Component Analysis: PCA）では，行列要素は負値をとることが許されている。従って，式 (2) のように入力を基底の線形和で表現する際に，基底の加減算による細かな調節が可能となる。一方，NMF では，基底の加算しか許されず，いったんアクティベートされた基底の影響を打ち消すことはできない。そのため，各基底 $\boldsymbol{w}_k$ が $\boldsymbol{x}_n$ 中の局所的な「パーツ」に対応し，少数の基底で $\boldsymbol{x}_n$ を表現する方が都合が良い。

このようなパーツに基づく分解表現は，音楽音響信号の分解と相性が良い。なぜなら，調波音のスペクトルは周波数軸上でスパースであり，混合音スペクトルは局所的な周波数帯域上のパーツの組み合わせと見なせるからである。

2.4 音源分離への応用

観測信号の複素スペクトログラムを $\tilde{\boldsymbol{X}} = [\tilde{\boldsymbol{x}}_1, \cdots, \tilde{\boldsymbol{x}}_N] \in \mathbb{C}^{M\times N}$，$k$ 番目の音源信号の複素スペクトログラムを $\tilde{\boldsymbol{X}}_k = [\tilde{\boldsymbol{x}}_{k1}, \cdots, \tilde{\boldsymbol{x}}_{kN}] \in \mathbb{C}^{M\times N}$ とする。M は周波数ビン数，N はフレーム数である。観測した混合音が K 個の音源信号の



図–3 パワースペクトログラムに対する IS ダイバージェンスに基づく非負値行列分解（IS-NMF）の適用結果。

瞬時混合であると仮定すると，以下が成立する。

$$\tilde{\boldsymbol{X}} = \sum_{k=1}^{K} \tilde{\boldsymbol{X}}_k \quad \left(\tilde{\boldsymbol{x}}_n = \sum_{k=1}^{K} \tilde{\boldsymbol{x}}_{kn} \right) \tag{5}$$

観測変数 $\tilde{\boldsymbol{X}}$ を潜在変数 $\tilde{\boldsymbol{X}}_k$ に分解する問題は不良設定であるので，$\tilde{\boldsymbol{X}}_k$ に対応するパワースペクトログラム $\boldsymbol{X}_k = [\boldsymbol{x}_{k1}, \cdots, \boldsymbol{x}_{kN}] \in \mathbb{R}_+^{M\times N}$ ($x_{knm} = |\tilde{x}_{knm}|^2$) は，ランク 1 行列 $\boldsymbol{Y}_k = [\boldsymbol{y}_{k1}, \cdots, \boldsymbol{y}_{kN}] \in \mathbb{R}_+^{M\times N}$ で近似する（図–3）。

$$\boldsymbol{X}_k \approx \boldsymbol{w}_k \boldsymbol{h}_k^{\mathrm{T}} \stackrel{\mathrm{def}}{=} \boldsymbol{Y}_k \tag{6}$$

すなわち，$\boldsymbol{Y}_k$ の任意のフレーム n におけるパワースペクトル $\boldsymbol{y}_{kn}$ は基底スペクトル $\boldsymbol{w}_k \in \mathbb{R}^M$ を重み h_{kn} でスケーリングするだけで得られるという仮定をおいた（$\boldsymbol{y}_{kn} = h_{kn}\boldsymbol{w}_k$）。

まず，潜在変数 $\tilde{\boldsymbol{x}}_{kn}$ が $\boldsymbol{y}_{kn}$ で定まる対角共分散行列を持つ複素ガウス分布に従うことを仮定する。

$$\tilde{\boldsymbol{x}}_{kn} \sim \mathcal{N}_c(\boldsymbol{0}, \mathrm{diag}(\boldsymbol{y}_{kn})) \tag{7}$$

ただし，$\mathrm{diag}(\boldsymbol{\eta})$ はベクトル $\boldsymbol{\eta}$ を対角成分に持つ対角行列を表す。式 (5) に着目すると，複素ガウス分布の再生性から

$$\tilde{\boldsymbol{x}}_n \sim \mathcal{N}_c(\boldsymbol{0}, \mathrm{diag}(\boldsymbol{y}_n)) \tag{8}$$

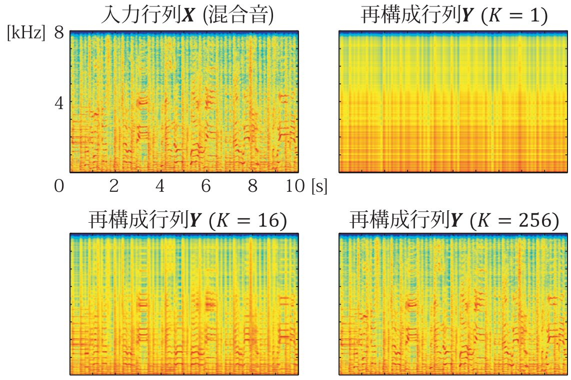
を得る。ただし，$\boldsymbol{y}_n = \sum_k \boldsymbol{y}_{kn}$ である。従って，$x_{nm} = |\tilde{x}_{nm}|^2$ は指数分布に従うことが分かる。

$$x_{nm} \sim \mathrm{Exponential}(y_{nm}) \tag{9}$$

ここで，式 (8) の対数をとって符号反転させると，式 (4) と定数項を除いて等しい。従って，式 (8) の最大化（最尤推定）は式 (4) の最小化と等価であり，IS-NMF の適用が適切であると分かる。

最終的に，式 (7) 及び式 (8) に着目すると，$\tilde{\boldsymbol{x}}_n$ が与えられたときの $\tilde{\boldsymbol{x}}_{kn}$ の事後分布は複素ガウス分布になることが分かり，その平均と分散は

$$\mathbb{E}[\tilde{\boldsymbol{x}}_{kn}|\tilde{\boldsymbol{x}}_n] = \mathrm{diag}(\boldsymbol{y}_{kn})\mathrm{diag}(\boldsymbol{y}_n)^{-1}\tilde{\boldsymbol{x}}_n \tag{10}$$



図–4 異なる基底数 K に対する IS-NMF の結果。

$$\mathbb{V}[\tilde{\boldsymbol{x}}_{kn}|\tilde{\boldsymbol{x}}_n] = \mathrm{diag}(\boldsymbol{y}_{kn}) - \mathrm{diag}(\boldsymbol{y}_{kn})\mathrm{diag}(\boldsymbol{y}_n)^{-1}\mathrm{diag}(\boldsymbol{y}_{kn}) \tag{11}$$

で与えられる。この処理はウィナーフィルタリングと呼ばれ，$\tilde{\boldsymbol{X}}_k$ の位相は $\tilde{\boldsymbol{X}}$ の位相と同一であるという仮定が置かれている。最後に，逆フーリエ変換を用いて，$\mathbb{E}[\tilde{\boldsymbol{X}}_k|\tilde{\boldsymbol{X}}]$ から k 番目の音源信号を復元することができる。

基底数 K を変化させながら IS-NMF を適用した結果を図–4 に示す。K が小さすぎると近似が荒く，K を大きくしすぎると物理的な意味を持たない極めて局所的な基底ばかりになる。このことから，基底数 K を適切に定める重要性が分かる。

実際には，性質の良くない局所解に陥り易い IS-NMF の代わりに KL-NMF が利用される場合が多い。このとき，$\boldsymbol{X}$ や $\boldsymbol{X}_k$ は振幅スペクトログラムとするのが一般的である（$x_{knm} = |\tilde{x}_{knm}|$）。

まず，潜在変数 x_{knm} が y_{knm} で定まるポアソン分布に従うことを仮定する。

$$x_{knm} \sim \mathrm{Poisson}(y_{knm}) \tag{12}$$

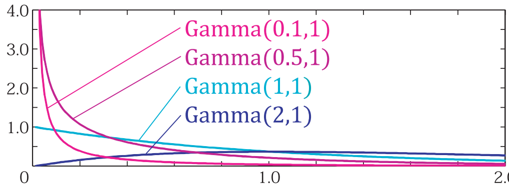
ここで，複数の音源信号の重畳における振幅スペクトルの加法性を仮定すると（実際には成立しないことに注意），ポアソン分布の再生性から

$$x_{nm} \sim \mathrm{Poisson}(y_{nm}) \tag{13}$$

を得る。ここで，式 (13) の対数をとって符号反転させると，式 (3) と定数項を除いて等しい。従って，式 (13) の最大化（最尤推定）は式 (3) の最小化と等価である。

2.5 ノンパラメトリックベイズモデル

2.3 節で述べたように，NMF は最尤推定であってもスパースな解が得られ易いが，適切な事前分布を導入してベイズ推定を行うことで，よりスパースな解を得ることができる。更に，ノンパラメトリッ



図–5　形状母数が異なる幾つかのガンマ分布。

クベイズモデルを定式化すれば，基底数を $K \to \infty$ とした場合でもスパースな学習が可能になる。すなわち，観測行列 $\boldsymbol{X}$ に合わせて高々有限個の基底がアクティベートされるような機構が実現できる。具体的には，ガンマ過程あるいはベータ過程を事前分布に用いることになる。

2.5.1 ガンマ過程に基づく基底数の無限化

ガンマ過程に基づく NMF のノンパラメトリックベイズモデル（GaP-NMF）について説明する。まず，式 (2) に対し，K 次元の非負値ベクトル $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_K]$ を導入する。

$$\boldsymbol{x}_n \approx \sum_{k=1}^{K} \theta_k h_{kn} \boldsymbol{w}_k \stackrel{\text{def}}{=} \boldsymbol{y}_n \tag{14}$$

ここで，$\theta_k \geq 0$ は基底 k の大域的な重みである。この $\boldsymbol{\theta}$ に対し，観測データ $\boldsymbol{X}$ を表現するのに必要な基底 k 以外の要素 θ_k がゼロとなるようなスパースな学習を行いたい。

ノンパラメトリックベイズモデルを定式化するため，$\boldsymbol{\theta}$，$\boldsymbol{W}$，$\boldsymbol{H}$ に対して事前分布を導入する。まず，$\boldsymbol{W}$ 及び $\boldsymbol{H}$ の各要素は非負値であるので，ガンマ事前分布を用いると都合が良い。

$$w_{km} \sim \mathrm{Gamma}(a_0^w, b_0^w) \tag{15}$$

$$h_{kn} \sim \mathrm{Gamma}(a_0^h, b_0^h) \tag{16}$$

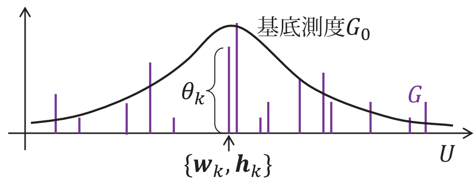
ここで，$a_0^* > 0$ 及び $b_0^* > 0$ はそれぞれ，ガンマ分布の形状母数と逆尺度母数である。更に，$\boldsymbol{\theta}$ に対しても同様にガンマ事前分布を仮定する。

$$\theta_k \sim \mathrm{Gamma}\left(\frac{\alpha c}{K}, \alpha\right) \tag{17}$$

ここで，$\alpha > 0$ 及び $c > 0$ は超パラメータである。ガンマ分布の形状母数が小さくなるほど 0 が出る確率が大きくなる（図–5）。ただし，$\mathbb{E}_{\mathrm{prior}}[\theta_k] = \frac{c}{K}$，$\mathbb{E}_{\mathrm{prior}}[\sum_k \theta_k] = c$ である。

ここで，式 (15)，式 (16) 及び式 (17) で構成される有限モデルに対して，$K \to \infty$ となる極限を考えると，以下のガンマ過程が得られる。

$$G \sim \mathrm{GaP}(\alpha, G_0) \tag{18}$$



図–6　NMF のためのガンマ過程事前分布。

ここで，G_0 は空間 U（$\boldsymbol{w} \in \mathbb{R}_+^M$ と $\boldsymbol{h} \in \mathbb{R}_+^N$ の直積空間）上に定義された基底測度であり，$G_0(U) = c$ を満たす（図–6）。このとき，G は U 上の離散測度となり，空間 U の任意の分割 $\{U_i\}_{i=1}^{I}$ に対して

$$G(U_i) \sim \mathrm{Gamma}(\alpha G_0(U_i), \alpha) \tag{19}$$

が成立している。ただし，$\mathbb{E}[G] = G_0$ である。微小区間への分割を $\{U_k\}_{k=1}^{\infty}$ とすると，$G(U_k) = \theta_k$ である。α は集中度と呼ばれ，α が小さくなるほど $\boldsymbol{\theta}$ はよりスパースになる。計算機上では $K \to \infty$ は扱えないが，K を α に比べて十分大きな値に設定すれば，式 (17) はガンマ過程の良い近似となる（weak-limit approximation）。

2.5.2 変分ベイズ学習

前項の議論を踏まえて，式 (13), (15), (16), (17) で定義されるノンパラメトリックベイズ KL-NMF（GaP-KL-NMF）あるいは式 (8), (15), (16), (17) で定義されるノンパラメトリックベイズ IS-NMF（GaP-IS-NMF）に対する変分ベイズ法（Variational Bayes: VB）について述べる。今，観測データ $\boldsymbol{X}$ が与えられたときに，ベイズの定理を用いて未知パラメータ $\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}$ の事後分布

$$p(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H} | \boldsymbol{X}) = \frac{p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})}{p(\boldsymbol{X})} \tag{20}$$

を計算したい。しかし，周辺尤度 $p(\boldsymbol{X})$ は解析的に計算できないため，対数周辺尤度の変分下限 $\mathcal{L}(q)$ を構成し，逐次最大化を行うことで $p(\boldsymbol{X})$ を近似することを考える。すなわち，変分分布 $q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})$ を定義し，凹関数 $f(x) = \log(x)$ に対して Jensen の不等式を用いると以下を得る。

$$\begin{aligned}
&\log p(\boldsymbol{X}) \\
&= \log \int q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) \frac{p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})}{q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})} d\boldsymbol{\theta} d\boldsymbol{W} d\boldsymbol{H} \\
&\geq \int q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) \log \frac{p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})}{q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})} d\boldsymbol{\theta} d\boldsymbol{W} d\boldsymbol{H} \\
&= \mathbb{E}_{q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})}[\log p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})] \\
&\quad - \mathbb{E}_{q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})}[\log q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})] \\
&\stackrel{\text{def}}{=} \mathcal{L}(q)
\end{aligned} \tag{21}$$

Algorithm 3 GaP-KL-NMF のベイズ推定

Require: 非負値行列 $\boldsymbol{X} \in \mathbb{R}_+^{M\times N}$, 最大基底数 K, ガンマ過程の集中度 α, ガンマ分布のパラメータ $a_0^w, b_0^w, a_0^h, b_0^h$

変分事後分布 $q(\boldsymbol{\theta})$, $q(\boldsymbol{W})$, $q(\boldsymbol{H})$ をランダムに初期化
while not converged **do**
　$\lambda_{knm} \propto \mathbb{E}_q[\theta_k w_{km} h_{kn}]$
　$q(\theta_k) = \mathrm{Gamma}(\frac{\alpha c}{K} + \sum_{nm} \lambda_{knm} x_{nm},$
　　　$\alpha + \sum_{nm} \mathbb{E}_q[w_{km} h_{kn}])$
　$q(w_{km}) = \mathrm{Gamma}(a_0^w + \sum_n \lambda_{knm} x_{nm},$
　　　$b_0^w + \sum_n \mathbb{E}_q[\theta_k h_{kn}])$
　$q(h_{kn}) = \mathrm{Gamma}(a_0^h + \sum_m \lambda_{knm} x_{nm},$
　　　$b_0^h + \sum_m \mathbb{E}_q[\theta_k w_{km}])$
end while
return 変分事後分布 $q(\boldsymbol{\theta})$, $q(\boldsymbol{W})$, $q(\boldsymbol{H})$

等号成立条件は $q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) = p(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}|\boldsymbol{X})$ であり，このとき $\mathcal{L}(q)$ が最大値をとる。しかし，真の事後分布 $p(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}|\boldsymbol{X})$ は計算困難であるため，変分事後分布を因子分解可能な形 $q(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) = q(\boldsymbol{\theta})q(\boldsymbol{W})q(\boldsymbol{H})$ に限定し，その中でも以下で計算できる変分下限

$$\begin{aligned}\mathcal{L}(q) = &\mathbb{E}_q[\log p(\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})] + \mathbb{E}_q[\log p(\boldsymbol{\theta})] \\ &+ \mathbb{E}_q[\log p(\boldsymbol{W})] + \mathbb{E}_q[\log p(\boldsymbol{H})] \\ &+ H(q(\boldsymbol{\theta})) + H(q(\boldsymbol{W})) + H(q(\boldsymbol{H}))\end{aligned} \tag{22}$$

を最大化するものを求めたい。ここで，$H(\cdot)$ はエントロピーを表す。これは，変分事後分布 $q(\boldsymbol{\theta})q(\boldsymbol{W})q(\boldsymbol{H})$ の真の事後分布 $p(\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}|\boldsymbol{X})$ に対する KL ダイバージェンスを最小化することと等価である。式 (22) を逐次最大化するには，以下の更新式を収束するまで繰り返せば良い。

$$q(\boldsymbol{\theta}) \propto \exp(\mathbb{E}_{q(\boldsymbol{H}, \boldsymbol{W})}[\log p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})]) \tag{23}$$

$$q(\boldsymbol{H}) \propto \exp(\mathbb{E}_{q(\boldsymbol{\theta}, \boldsymbol{W})}[\log p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})]) \tag{24}$$

$$q(\boldsymbol{W}) \propto \exp(\mathbb{E}_{q(\boldsymbol{\theta}, \boldsymbol{H})}[\log p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})]) \tag{25}$$

2.5.3 GaP-KL-NMF

まず，GaP-KL-NMF に対する VB を導出する。式 (22) で与えられる変分下限 $\mathcal{L}(q)$ の第 1 項は式 (13) で計算できる対数ポアソン尤度の期待値であるが，依然として解析的に計算できない。そのため，凹関数 $f(x) = \log(x)$ に対して Jensen の不等式を用いると更なる変分下限

Algorithm 4 GaP-IS-NMF のベイズ推定

Require: 非負値行列 $\boldsymbol{X} \in \mathbb{R}_+^{M\times N}$, 最大基底数 K, ガンマ過程の集中度 α, ガンマ分布のパラメータ $a_0^w, b_0^w, a_0^h, b_0^h$

変分事後分布 $q(\boldsymbol{\theta})$, $q(\boldsymbol{W})$, $q(\boldsymbol{H})$ をランダムに初期化
while not converged **do**
　$\lambda_{knm} \propto \mathbb{E}_q[\theta_k^{-1} w_{km}^{-1} h_{kn}^{-1}]^{-1}$
　$\omega_{nm} \propto \sum_k \mathbb{E}_q[\theta_k w_{km} h_{kn}]$
　$q(\theta_k) = \mathrm{GIG}(\frac{\alpha c}{K}, \alpha + \sum_{nm} \omega_{nm}^{-1} \mathbb{E}_q[w_{km}]\mathbb{E}_q[h_{kn}],$
　　　$\sum_{nm} x_{nm} \lambda_{knm}^2 \mathbb{E}_q[w_{mk}^{-1}]\mathbb{E}_q[h_{kn}^{-1}])$
　$q(w_{km}) = \mathrm{GIG}(a_0^w, b_0^w + \sum_n \omega_{nm}^{-1} \mathbb{E}_q[\theta_k]\mathbb{E}_q[h_{kn}],$
　　　$\sum_n x_{nm} \lambda_{knm}^2 \mathbb{E}_q[\theta_k^{-1}]\mathbb{E}_q[h_{kn}^{-1}])$
　$q(h_{kn}) = \mathrm{GIG}(a_0^h, \sum_m \omega_{nm}^{-1} \mathbb{E}_q[\theta_k]\mathbb{E}_q[w_{km}],$
　　　$\sum_m x_{nm} \lambda_{knm}^2 \mathbb{E}_q[\theta_k^{-1}]\mathbb{E}_q[w_{km}^{-1}])$
end while
return 変分事後分布 $q(\boldsymbol{\theta})$, $q(\boldsymbol{W})$, $q(\boldsymbol{H})$

$$\begin{aligned}&\mathbb{E}_q[\log p(\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})] \\ &\overset{c}{=} \mathbb{E}_q\left[\sum_{nm}\left(x_{nm}\log\sum_k y_{knm} - \sum_k y_{knm}\right)\right] \\ &= \sum_{nm} x_{nm}\mathbb{E}_q\left[\log\sum_k \lambda_{knm}\frac{y_{knm}}{\lambda_{knm}}\right] \\ &\quad - \sum_{knm}\mathbb{E}_q[y_{knm}] \\ &\geq \sum_{nm} x_{nm}\sum_k \lambda_{knm}\mathbb{E}_q\left[\log\frac{y_{knm}}{\lambda_{knm}}\right] \\ &\quad - \sum_{knm}\mathbb{E}_q[y_{knm}] \\ &\overset{\mathrm{def}}{=} \mathbb{E}_q[\log q(\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})]\end{aligned} \tag{26}$$

を得る。ここで，λ_{knm} は $\sum_k \lambda_{knm} = 1$ を満たす補助変数である。等号成立条件（変分下限が最大となる条件）はラグランジュの未定乗数法を用いて求めることができ，$\lambda_{knm} \propto \mathbb{E}_q[y_{knm}]$ となる。

最後に，各パラメータに対する変分事後分布を導出する。実際には式 (22) で与えられる元の変分下限 $\mathcal{L}(q)$ ではなく，式 (26) を用いて得られた更なる変分下限を最大化することになる。その結果，式 (23), (24), (25) において，$\log p(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H})$ の代わりに次式を用いることになる。

$$\begin{aligned}\log q(\boldsymbol{X}, \boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) &= \log q(\boldsymbol{X}|\boldsymbol{\theta}, \boldsymbol{W}, \boldsymbol{H}) \\ &+ \log p(\boldsymbol{\theta}) + \log p(\boldsymbol{W}) + \log p(\boldsymbol{H})\end{aligned} \tag{27}$$

具体的には，最適な変分事後分布 $q(\boldsymbol{\theta})$ は，$\boldsymbol{\theta}$ に関連する項のみを取り出すと以下のとおり計算できる。

$$
\begin{aligned}
&\log q(\boldsymbol{\theta}) \\
&\quad \overset{c}{=} \sum_{knm} x_{nm}\lambda_{knm}\log\theta_k \\
&\quad - \sum_{knm}\theta_k \mathbb{E}_q[w_{km}]\,\mathbb{E}_q[h_{kn}] \\
&\quad + \sum_k \left(\left(\frac{\alpha c}{K}-1\right)\log\theta_k - \alpha\theta_k\right) \qquad (28)
\end{aligned}
$$

従って，θ_k の事後分布はガンマ分布となる。同様に，最適な $q(\boldsymbol{W})$ や $q(\boldsymbol{H})$ もガンマ分布として求まる。**Algorithm 3** に更新則を示す。反復ごとに，$\mathbb{E}[\theta_k]$ が十分に小さい基底 k を削除していけば，実効的な基底数 K_+ が自動的に定まる。

2.5.4 GaP-IS-NMF

次に，GaP-IS-NMF に対する VB [3] を導出する。式 (22) で与えられる変分下限 $\mathcal{L}(q)$ の第 1 項は式 (9) で与えられる対数指数尤度の期待値であり，やはり解析的に計算できない。そこで，凹関数 $f(x)=-\frac{1}{x}$ に対して Jensen の不等式を考える。

$$
-\frac{1}{\sum_{k=1}^K x_k} = -\frac{1}{\sum_{k=1}^K \lambda_k \frac{x_k}{\lambda_k}} \geq \sum_{k=1}^K \frac{\lambda_k^2}{x_k} \qquad (29)
$$

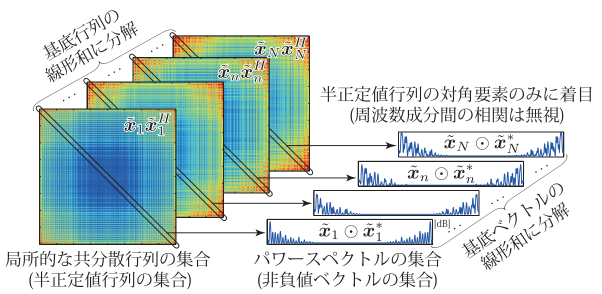
ここで，λ_k は $\sum_k \lambda_k = 1$ を満たす補助変数であり，等号成立条件は $\lambda_k \propto x_k$ である。更に，凸関数 $g(x)=-\log(x)$ に対する 1 次のテイラー展開（ω における接線）を考える。

$$
-\log(x) \geq -\log(\omega) - \frac{x}{\omega} + 1 \qquad (30)
$$

ここで，ω は補助変数であり，等号成立条件は $\omega = x$ である。これら二つの不等式を用いると変分下限 $\mathbb{E}_q[\log q(\boldsymbol{X}|\boldsymbol{\theta},\boldsymbol{W},\boldsymbol{H})]$ を得る。

$$
\begin{aligned}
&\mathbb{E}_q[\log p(\boldsymbol{X}|\boldsymbol{\theta},\boldsymbol{W},\boldsymbol{H})] \\
&\quad \overset{c}{=} \mathbb{E}_q\left[\sum_{nm}\left(-x_{nm}\,(y_{nm})^{-1} - \log\,(y_{nm})\right)\right] \\
&\quad \geq -\sum_{nm} x_{nm}\mathbb{E}_q\left[\sum_k \frac{\lambda_{knm}^2}{y_{knm}}\right] \\
&\quad - \sum_{nm}\left(\log(\omega_{nm}) + \mathbb{E}_q\left[\frac{y_{nm}}{\omega_{nm}}\right] - 1\right) \\
&\quad \overset{\text{def}}{=} \mathbb{E}_q[\log q(\boldsymbol{X}|\boldsymbol{\theta},\boldsymbol{W},\boldsymbol{H})] \qquad (31)
\end{aligned}
$$

各パラメータに対する変分事後分布は，2.5.3 項と同様に求められる。具体的には，最適な変分事後分布 $q(\boldsymbol{\theta})$ は，$\boldsymbol{\theta}$ に関連する項のみを取り出すと



図–7　音源分離のための半正定値テンソル分解（PSDTF）。

$$
\begin{aligned}
&\log q(\boldsymbol{\theta}) \\
&\quad \overset{c}{=} -\sum_{knm} x_{nm}\lambda_{knm}^2\theta_k^{-1}\mathbb{E}_q\big[w_{km}^{-1}\big]\,\mathbb{E}_q\big[h_{kn}^{-1}\big] \\
&\quad - \sum_{knm}\omega_{nm}^{-1}\theta_k\mathbb{E}_q[w_{km}]\,\mathbb{E}_q[h_{kn}] \\
&\quad + \sum_k \left(\left(\frac{\alpha c}{K}-1\right)\log\theta_k - \alpha\theta_k\right) \qquad (32)
\end{aligned}
$$

従って，θ_k の事後分布は Generalized Inverse Gaussian（GIG）分布となることが分かる（詳細は [3] 参照）。**Algorithm 4** に更新則を示す。

3. 半正定値テンソル分解

本章では，NMF の自然な拡張である半正定値テンソル分解 [4, 5]（PSDTF）について解説する。PSDTF では，各フレーム n の複素スペクトル $\tilde{\boldsymbol{x}}_n$ の自己共分散 $\boldsymbol{X}_n = \tilde{\boldsymbol{x}}_n\tilde{\boldsymbol{x}}_n^H$，すなわち**半正定値行列を少数の半正定値行列の和に分解**する（図–7）。一方，NMF では，上記行列の対角成分（パワースペクトル）$\boldsymbol{x}_n = \tilde{\boldsymbol{x}}_n \odot \tilde{\boldsymbol{x}}_n^*$，すなわち**非負値ベクトルを少数の非負値ベクトルの和に分解**する。行列の半正定値性はベクトルの非負値性の拡張概念であり，非負値テンソル分解（Nonnegative Tensor Factorization: NTF）と PSDTF とは異なる。

音源分離においては，観測スペクトログラム $\tilde{\boldsymbol{X}}$ から，式 (5) を満たす音源スペクトログラム $\tilde{\boldsymbol{X}}_k$ の位相を推定できる。音源信号の周期と短時間フーリエ変換の窓長 M が異なると，音源信号の巡回定常性の仮定が成り立たないため，周波数ビン間の相関を取り扱える利点は大きい。マルチチャネル音源分離において，マイク間の相関行列を分解する際にも同様のモデルが提案されている [10]。

3.1 コスト関数最小化としての定式化

PSDTF では，観測データとして 3 階のテンソル $\boldsymbol{X} = [\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n] \in \mathbb{C}^{M\times M\times N}$ に対する分解を行う。各要素 $\boldsymbol{X}_n \succeq \boldsymbol{0} \in \mathbb{C}^{M\times M}$ は半正定値

行列とする。今，各 $\boldsymbol{X}_n$ を K 個の半正定値行列 $\{\boldsymbol{W}_k\}_{k=1}^K$（基底行列）の凸結合で近似したい。

$$\boldsymbol{X}_n \approx \sum_{k=1}^{K} h_{kn}\boldsymbol{W}_k \overset{\text{def}}{=} \boldsymbol{Y}_n \tag{33}$$

ここで，$h_{kn} \geq 0$ は $\boldsymbol{X}_n$ における基底行列 $\boldsymbol{W}_k$ の重みである。観測行列 $\boldsymbol{X}_n$ と再構成行列 $\boldsymbol{Y}_n$ との間の誤差 $\mathcal{D}(\boldsymbol{X}_n|\boldsymbol{Y}_n)$ を評価する尺度として，非負値ベクトル間の KL ダイバージェンスや IS ダイバージェンスの拡張である，半正定値行列間の von-Neumann（vN）ダイバージェンスや Log-Determinant（LD）ダイバージェンスがある [11]。

$$\begin{aligned}\mathcal{D}_{\text{vN}}(\boldsymbol{X}_n|\boldsymbol{Y}_n) =\ & \mathrm{tr}(\boldsymbol{X}_n \log \boldsymbol{X}_n - \boldsymbol{X}_n \log \boldsymbol{Y}_n \\ & -\boldsymbol{X}_n + \boldsymbol{Y}_n)\end{aligned} \tag{34}$$

$$\begin{aligned}\mathcal{D}_{\text{LD}}(\boldsymbol{X}_n|\boldsymbol{Y}_n) =\ & \mathrm{tr}(\boldsymbol{X}_n\boldsymbol{Y}_n^{-1}) \\ & -\log\left|\boldsymbol{X}_n\boldsymbol{Y}_n^{-1}\right| - M\end{aligned} \tag{35}$$

3.2 乗法更新アルゴリズムに基づく最適化

コスト関数 $\mathcal{D}(\boldsymbol{X}|\boldsymbol{Y}) = \sum_n \mathcal{D}(\boldsymbol{X}_n|\boldsymbol{Y}_n)$ を最小化する $\boldsymbol{H} = [\boldsymbol{h}_1, \cdots, \boldsymbol{h}_K] \in \mathbb{R}^{N\times K}$ 及び $\boldsymbol{W} = [\boldsymbol{W}_1, \cdots, \boldsymbol{W}_K] \in \mathbb{C}^{M\times M\times K}$ を求めるため, LD-PSDTF に対しても乗法更新アルゴリズム [4, 5] が提案されている。更新則は **Algorithm 5** で与えられる（導出は文献 [4, 5] 参照）。h_{kn} の非負性と $\boldsymbol{W}_k$ の半正定値性は自然に保たれているが, $\mathrm{tr}(\boldsymbol{W}_k) = 1$ を満たすよう，反復ごとに $\boldsymbol{W}_k$ 及び $\boldsymbol{h}_k$ をスケーリングしておく。

Algorithm 5 LD-PSDTF の最尤推定

Require: テンソル $\boldsymbol{X} \in \mathbb{C}^{M\times M\times N}$, 基底数 K
基底テンソル $\boldsymbol{W} \in \mathbb{C}^{M\times M\times K}$ をランダムに初期化
非負値行列 $\boldsymbol{H} \in \mathbb{R}_+^{M\times K}$ をランダムに初期化
while not converged **do**
　$\boldsymbol{P}_k = \sum_{n=1}^N h_{kn}\boldsymbol{Y}_n^{-1}$
　$\boldsymbol{Q}_k = \sum_{n=1}^N h_{kn}\boldsymbol{Y}_n^{-1}\boldsymbol{X}_n\boldsymbol{Y}_n^{-1}$
　コレスキー分解 $\boldsymbol{Q}_k = \boldsymbol{L}_k\boldsymbol{L}_k^{\mathrm{T}}$
　$\boldsymbol{W}_k \leftarrow \boldsymbol{W}_k\boldsymbol{L}_k(\boldsymbol{L}_k^{\mathrm{T}}\boldsymbol{W}_k\boldsymbol{P}_k\boldsymbol{W}_k\boldsymbol{L}_k)^{-\frac{1}{2}}\boldsymbol{L}_k^{\mathrm{T}}\boldsymbol{W}_k$
　$h_{kn} \leftarrow h_{kn}\left(\frac{\mathrm{tr}(\boldsymbol{Y}_n^{-1}\boldsymbol{W}_k\boldsymbol{Y}_n^{-1}\boldsymbol{X}_n)}{\mathrm{tr}(\boldsymbol{Y}_n^{-1}\boldsymbol{W}_k)}\right)^{\frac{1}{2}}$
end while
return 基底テンソル $\boldsymbol{W}$，非負値行列 $\boldsymbol{H}$

3.3 音源分離への応用

式 (5) を満たすように，$\tilde{\boldsymbol{x}}_n$ を $\{\tilde{\boldsymbol{x}}_{kn}\}_{k=1}^K$ の和に分解したい。まず，潜在変数 $\tilde{\boldsymbol{x}}_{kn}$ が共分散行列 $\boldsymbol{Y}_{kn}$ を持つ複素ガウス分布に従うことを仮定する。

$$\tilde{\boldsymbol{x}}_{kn} \sim \mathcal{N}_c(\boldsymbol{0}, \boldsymbol{Y}_{kn}) \tag{36}$$

ここで，式 (7) のように共分散行列を対角行列に限定しないことで，周波数ビン間の相関を考慮している。式 (5) と複素ガウス分布の再生性から

$$\tilde{\boldsymbol{x}}_n \sim \mathcal{N}_c(\boldsymbol{0}, \boldsymbol{Y}_n) \tag{37}$$

を得る。ただし，$\boldsymbol{Y}_n = \sum_k \boldsymbol{Y}_{kn}$ である。ここで，式 (37) の対数をとって符号反転させると，式 (35) と定数項を除いて等しい。従って，式 (37) の最大化は式 (35) の最小化と等価であり，LD-PSDTF を用いて $\boldsymbol{Y}_n$ や $\boldsymbol{Y}_{kn}$ を求めることができる。

最終的に，式 (36), (37) から，$\tilde{\boldsymbol{x}}_n$ が与えられたときの $\tilde{\boldsymbol{x}}_{kn}$ の事後分布は複素ガウス分布になることが分かり，その平均と分散は次式となる。

$$\mathbb{E}[\tilde{\boldsymbol{x}}_{kn}|\tilde{\boldsymbol{x}}_n] = \boldsymbol{Y}_{kn}\boldsymbol{Y}_n^{-1}\tilde{\boldsymbol{x}}_n \tag{38}$$

$$\mathbb{V}[\tilde{\boldsymbol{x}}_{kn}|\tilde{\boldsymbol{x}}_n] = \boldsymbol{Y}_{kn} - \boldsymbol{Y}_{kn}\boldsymbol{Y}_n^{-1}\boldsymbol{Y}_{kn} \tag{39}$$

ここで，式 (10) とは異なり，$\tilde{\boldsymbol{X}}_k$ の位相は $\tilde{\boldsymbol{X}}$ の位相とは異なる点に注意する。IS-NMF のように各周波数ビン n, m ごとではなく，各フレーム n ごとに一挙に分離を行うことで，周波数ビン間の相関を考慮しながら高品質な分離が可能となる。

3.4 ベイズ拡張・今後の課題

文献 [4, 5] において，NMF と同様にガンマ過程を用いて基底数を $K \to \infty$ としたノンパラメトリックベイズモデルが提案されている。今後の課題として，vN-PSDTF に対する乗法更新則の導出や計算量の削減が挙げられる。

4. 歌声・伴奏音・打楽器音の分離

最後に，スパース性に基づいて音楽音響信号を楽器種別に分離する技術を紹介する。ロバスト主成分分析（RPCA）は，入力行列 $\boldsymbol{X}$ を低ランク行列 $\boldsymbol{L}$ とスパース行列 $\boldsymbol{S}$ の和に分解する。具体的には，$\boldsymbol{X} = \boldsymbol{L} + \boldsymbol{S}$ を満たすという制約のもとで，以下で定義される最適化問題を解く。

$$\text{minimize } \|\boldsymbol{L}\|_* + \lambda\|\boldsymbol{S}\|_1 \tag{40}$$

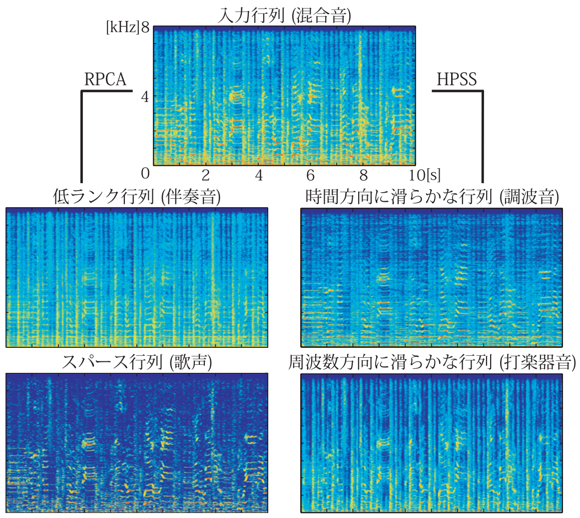
ここで，$\|\cdot\|_*$，$\|\cdot\|_1$ はそれぞれ核ノルム，L1 ノルムである。厳密ではないが拡張ラグランジュ法に基づく効率的な解法を **Algorithm 6** に示す。1 章で議論したとおり，混合音スペクトログラムを $\boldsymbol{X}$ として与えると，伴奏音スペクトログラム $\boldsymbol{L}$ と歌声スペクトログラム $\boldsymbol{S}$ が得られる。また，メディアンフィルタを時間方向・周波数方向に適用することで調波音・打楽器音を分離する

Algorithm 6 Inexact ALM に基づく RPCA

Require: 入力行列 $\boldsymbol{X} \in \mathbb{R}^{M \times N}$，重み係数 λ

初期化：$\boldsymbol{Y} = \boldsymbol{X} / \max(\|\boldsymbol{X}\|_2, \lambda^{-1}\|\boldsymbol{X}\|_\infty) \in \mathbb{R}^{M \times N}$
初期化：$\boldsymbol{S} = \boldsymbol{0} \in \mathbb{R}^{M \times N}$
初期化：$\mu > 0$ (*e.g.*, $\mu = 1.25/\|\boldsymbol{X}\|_2$)
初期化：$\rho > 1$ (*e.g.*, $\rho = 1.5$)
while not converged **do**
  $[\boldsymbol{U}, \boldsymbol{\Sigma}, \boldsymbol{V}] = \mathrm{SVD}(\boldsymbol{X} - \boldsymbol{S} + \mu^{-1}\boldsymbol{Y})$
  $\boldsymbol{L} \leftarrow \boldsymbol{U}\mathcal{F}_{\mu^{-1}}(\boldsymbol{\Sigma})\boldsymbol{V}^{\mathrm{T}}$
  $\boldsymbol{S} \leftarrow \mathcal{F}_{\lambda\mu^{-1}}(\boldsymbol{X} - \boldsymbol{L} + \mu^{-1}\boldsymbol{Y})$
  $\boldsymbol{Y} \leftarrow \boldsymbol{Y} + \mu(\boldsymbol{X} - \boldsymbol{L} - \boldsymbol{S})$
  $\mu \leftarrow \rho\mu$
end while
return 低ランク行列 $\boldsymbol{L}$, スパース行列 $\boldsymbol{S} \in \mathbb{R}^{M \times N}$

$^*\mathcal{F}_\epsilon[x] = x - \epsilon$ (if $x > \epsilon$), $x + \epsilon$ (if $x < -\epsilon$), 0 (otherwise)



図–8　音楽音響信号に対する RPCA・HPSS の適用結果

こと（Harmonic/Percussive Source Separation: HPSS）も可能である [7]。これらの技術による分離結果を**図–8** に示す。

5. おわりに

本稿では，スパース性に基づく音楽音響信号分解技術として，非負値行列分解（NMF），半正定値テンソル分解（PSDTF），ロバスト主成分分析（RPCA），調波・打楽器分離音（HPSS）を紹介した。これらは音声情報処理分野の影響を受けながら，音楽情報処理分野で独自の発展を遂げている。例えば，NMF（音響モデル）はソース・フィルタ型の楽器音モデルや，楽譜の生成モデル（言語モデル）と統合する試みが進んでいる。本稿が読者の理解の一助になれば幸いである。

謝　辞

本研究の一部は，JST CREST「OngaCREST プロジェクト」及び JSPS 科研費 26700020, 24220006，26280089 の支援を受けた。

文　献

[1] D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, **401**, 788–791 (1999).
[2] C. Févotte, N. Bertin and J.-L. Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence: With application to music analysis," *Neural Comput.*, **21**, 793–830 (2009).
[3] M. Hoffman, D. Blei and P. Cook, "Bayesian nonparametric matrix factorization for recorded music," *ICML-10*, pp. 439–446 (2010).
[4] K. Yoshii, R. Tomioka, D. Mochihashi and M. Goto, "Infinite positive semidefinite tensor factorization for source separation of mixture signals," *ICML-13*, pp. 576–584 (2013).
[5] K. Yoshii, R. Tomioka, D. Mochihashi and M. Goto, "Beyond NMF: Time-domain audio source separation without phase reconstruction," *ISMIR 2013*, pp. 369–374 (2013).
[6] Z. Lin, M. Chen and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," *Math. Prog.* (2010).
[7] D. FitzGerald, "Harmonic/percussive separation using median filtering," *DAFx-10* (2010).
[8] P. Smaragdis and J.C. Brown, "Non-negative matrix factorization for polyphonic music transcription," *WASPAA 2003*, pp. 177–180 (2003).
[9] 亀岡弘和, "非負値行列因子分解の音響信号処理への応用," 音響学会誌, **68**, 559–565 (2012).
[10] H. Sawada, H. Kameoka, S. Araki and N. Ueda, "Multichannel extensions of non-negative matrix factorization with complex-valued data," *IEEE Trans. TASLP*, **21**, 971–982 (2013).
[11] B. Kulis, M. Sustik and I. Dhillon, "Low-rank kernel learning with Bregman matrix divergences," *J. Mach. Learn. Res.*, **10**, 341–376 (2009).

付　録

本稿中で使用した確率密度関数を掲載しておく。

$$\mathrm{Gamma}(x|a,b) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx) \tag{A.1}$$

$$\mathrm{Exponential}(x|\lambda) = \mathrm{Gamma}(x|1, 1/\lambda) \tag{A.2}$$

$$\mathrm{Poisson}(x|\lambda) = \frac{\lambda^x}{x!} \exp(-\lambda) \tag{A.3}$$

$$\mathcal{N}_c(\boldsymbol{x}|\boldsymbol{0}, \boldsymbol{\Sigma}) = \frac{1}{\pi^M |\boldsymbol{\Sigma}|} \exp(-\boldsymbol{x}^H \boldsymbol{\Sigma}^{-1} \boldsymbol{x}) \tag{A.4}$$